

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-318766

(43)Date of publication of application : 16.11.2001

(51)Int.CI.

G06F 3/06  
G06F 12/08  
G06F 12/16  
G11B 19/02

(21)Application number : 2000-135013

(71)Applicant : NEC SOFTWARE SHIKOKU LTD

(22)Date of filing : 08.05.2000

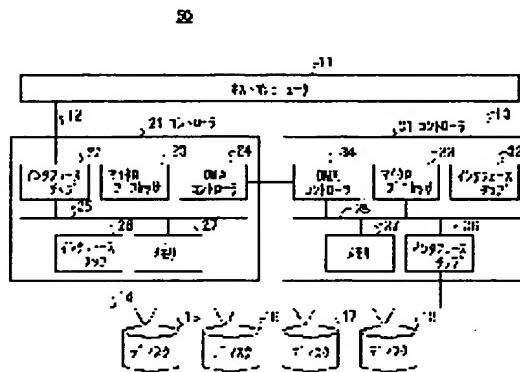
(72)Inventor : SHIRAIKI KAZUYA

## (54) DISK ARRAY DEVICE AND CACHE MEMORY CONTROL METHOD

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a disk array device and a cache memory control method for executing second and succeeding pieces of write command processing without putting loads on the processor of a main controller.

**SOLUTION:** This disk array device is composed of two controllers and executes the double write of write data in the two controllers without using the hardware of a shared memory. The write data are held in both of the two controllers and a means is provided for making the write data held in the other one of the two controllers mutually referable when a fault is generated in one of the controllers.



12. 15. 16...等のバッファ  
60...等のセクタレイアウト  
25...等の初期化

### LEGAL STATUS

[Date of request for examination] 17.04.2001

[Date of sending the examiner's decision of rejection] 12.02.2003

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

**Best Available Copy**

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office





ディスクアレイ装置およびキャッシュメモリ制御方法

特開2001-318766

実施の形態の構成に加えて、コントローラ1、31を増やすこともできる構造を備えている点に特徴を有している。具体的には、シーケンサードを送信するための、例えば、SCSIバスのような汎用バスと、自身のメモリ27、37のデータを任意のコントローラ1、31のメモリ27、37にダイレクトに送付できるDAMAコントローラ(不図示)を有している。

【0052】処理の流れは、上記第1の実施の形態の2コントローラ構成の場合と同じに同じで、2進路を走らし合うとするコントローラ1、31に対して毎回はシーケンサードを送信してゲートをシーザし、毎回はシーケンサードを無して直書きを行する。

【0053】なお、本発明が上記実施の形態に限定されず、本発明の技術思想の範囲内において、各部の形態は適宜変更可能であることは明らかである。また上記構成の数、位置、状態等は上記実施の形態に限定されず、本発明を実現する上で任意の数、位置、状態等をすることができる。また、各図において、同一構成要素には同一符号を付している。

【0054】【発明の効果】本発明は以上のように構成されているので、メモリコントローラのゲートをシーザしてそのキャッシュアドレスとビットマップを保持することにより、2日目以降のライトコマンド処理においてはメモリコントローラのプロセッサに負担をかけることなく実行できることになり、装置全体の性能の向上を図ることができるようになるといった効果を有する。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態に係るディスクアレ

ディスクアレイ装置およびキャッシュメモリ制御方法

特開2001-318766

イ装置を説明するための機能ブロック図である。【図2】図1のメモリに記載する構成の正真図である。【図3】本発明の第1の実施の形態に係るキャッシュメモリ制御方法におけるライト処理を説明するためのフローチャートである。

【図4】本発明の第1の実施の形態に係るディスクアレイ装置の各コントローラにおけるライト処理を説明するためのソローチャートである。

【図5】第1実施技術のキャッシュメモリ制御方式を説明するための機能ブロック図である。

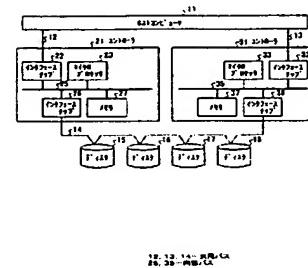
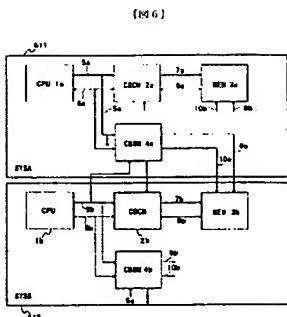
【図6】第2実施技術の記憶装置の制御方式を説明するための機能ブロック図である。

【図7】第3実施技術のディスクアレイ装置を説明するための機能ブロック図である。

【図2】

【図3】

【図4】



【図5】

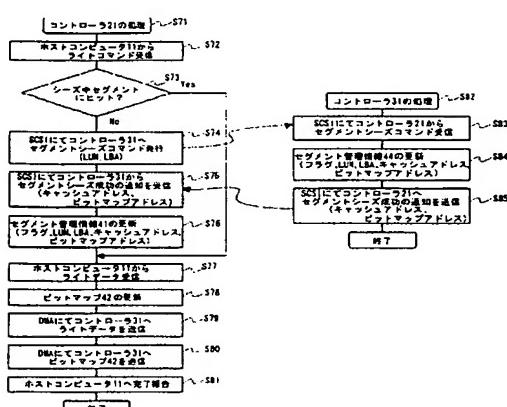
ディスクアレイ装置およびキャッシュメモリ制御方法

ディスクアレイ装置およびキャッシュメモリ制御方法

特開2001-318766

2001.04.24 10:41

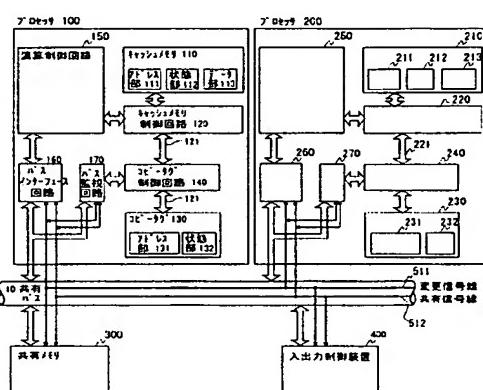
【図4】



ディスクアレイ装置およびキャッシュメモリ制御方法

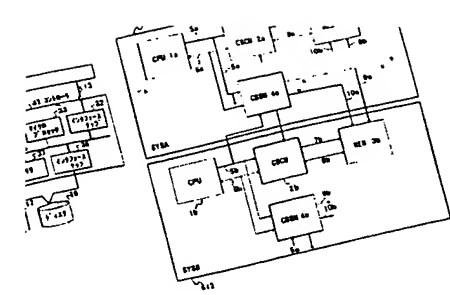
特開2001-318766

【図5】

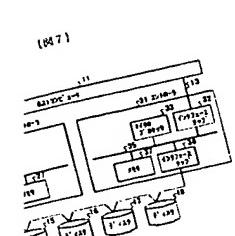


フロントページの続き

(S1) (a), Cl.	機別記号	F 1	F-7-F (参考)
G 0 6 F	I2/16	3 1 0	3 1 0 J
		3 2 0	3 2 0 I
G 1 1 B	I9/02	5 0 1	5 0 1 F



1047



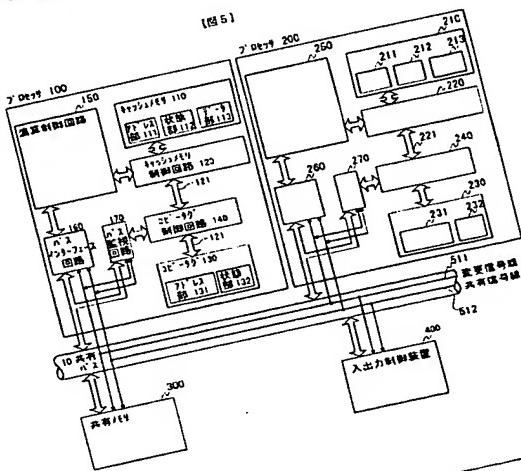
1047  
1048-1049

2003 04 24 10:41

- 10 -

ディスクアレイ装置およびキャッシュメモリ制御方法  
特開2001-318766

(55)



# IBM Total Storage Enterprise Server Model 800

## IBM Redbooks. SG24-6424-01. Second Edition (October 2002) ISBN (cont.)

## ISBN 0738426256 pp 88 ~ 89.

- Data that is accessed normally with some locality of reference will use partial track mode staging. This is the default mode.
- Data that is not a regular format, or where the history of access indicates that a full stage is required, will set the full track mode.
- The adaptive caching mode data is stored on disk and is reloaded at IMI.

### Sequential reads

Cache space is released according to Least Recently Used (LRU) algorithms. Space in the cache used for sequential data is freed up quicker than other cache or record data. The ESS will continue to pre-stage sequential tracks when the last few tracks in a sequential staging group are accessed.

Stage requests for sequential operations can be performed in parallel on the RAID array, giving the ESS its high sequential throughput characteristic. Parallel operations can take place because the logical data tracks are striped across the physical data disks in the RAID array.

## 3.28 NVS and write operations

### 3.28.1 Write operations

- Data written to an ESS is almost 100% fast write hits. A fast write hit occurs when the write I/O operation completes as soon as the data is in the ESS cache and non-volatile storage (NVS). The benefit of this is very fast write operations.

#### Fast write

Data received by the host adapter is transferred first to the NVS and a copy held in the host adapter buffer. The host is notified that the I/O operation is complete as soon as the data is in NVS. The host adapter, once the NVS transfer is complete, then transfers the data to the cache.

The data remains in the cache and NVS until it is destaged. Destage is triggered by cache and NVS usage thresholds.

### 3.28.2 NVS

The NVS size is 2 GB (1 GB per cluster). The NVS is protected by a battery. The battery will power the NVS for up to 72 hours following a total power failure.

#### NVS LRU

NVS is managed by a Least Recently Used (LRU) algorithm. The ESS attempts to keep free space in the NVS by anticipatory destaging of tracks when the space used in NVS exceeds a threshold. In addition, if the ESS is idle for a period of time, an idle destage function will destage tracks until, after about 5 minutes, all tracks will be destaged.

Both cache and NVS operate on LRU lists. Typically space in the cache occupied by sequential data is released earlier than space occupied by data that is likely to be re-referenced. Sequential data in the NVS is destaged ahead of random data.

When destaging tracks, the ESS attempts to destage all the tracks that would make up a RAID stripe, minimizing the RAID-related activities in the SSA adapter.

#### NVS location

NVS for cluster 1 is located physically in I/O drawer of cluster 2, and vice versa. This ensures that we always have one good copy of data, should we have a failure in one cluster.

See 3.8, "Cluster operation: failover/failback" on page 60 for more information.

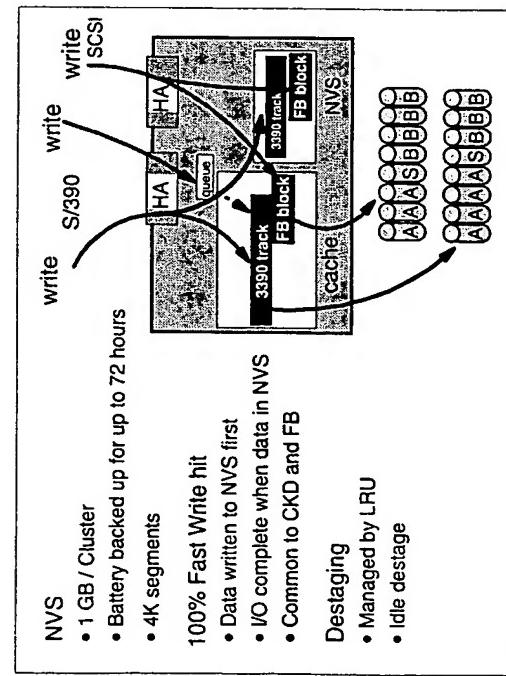


Figure 3-32 NVS - write

As Figure 3-32 illustrates, at any moment there are always two secured copies of any update into the ESS.

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.